**Method**

***Sample***

The survey was undertaken between March and September 2016. Census data and local informants converged in identifying nine villages in Kiunga where almost all West Papuans in the region lived. A preliminary house-to-house survey was conducted in the nine villages to confirm the composition and origins of all household residents. We defined West Papuans as native-born or the offspring of at least one parent from that territory. The full survey included 215 adolescents not included in the present analysis. Amongst the 643 adult residents, 157 (18.3%) were identified by others as long-term residents in the catchment area who were traveling elsewhere in PNG during the entire course of our study. The analytic sample therefore is based on 486 adult West Papuans. All those present in the catchment area agreed to participate, yielding a response rate of 100% of those with whom we made direct contact or 81.7% if adults who were absent were included in the calculation.

***Ethics***

Ethical permission for the study was provided by the University of New South Wales Human Research Ethics Committee and the Medical Research Council of PNG Ethics Committee.

***Measures***

The Refugee Mental Health Assessment Package (R-MHAP) is a comprehensive assessment tool assessing psychosocial factors and common mental disorders (CMDs) designed and developed for use amongst the West Papuan population (Tay et al., 2015b). Details of the qualitative and psychometric steps taken to develop, adapt and test the mental health measure have been provided previously (summarized in Supplementary File 1). In summary, after developing the draft modules, we undertook an extensive program of qualitative research to adapt the measures. This involved extensive consultations with the West Papuan community using focus groups and individual interviews. The aim was to capture idioms and terms that accurately reflected the constructs we were attempting to measure within the culture and context. Items of modules were modified to ensure their linguistic, semantic and contextual relevance. We then undertook extensive psychometric testing of the measures. In the final step, we conducted a concordance study in which we compared pooled cases on the R-MHAP (that is, persons who met diagnostic criteria for at least one diagnosis) with corresponding assignments on the SCID-5 (Tay et al., 2015b). The results indicated a sound level of agreement between the two measures: ROC=.9, sensitivity=.98, specificity=.97, positive

predictive power=.95, negative predictive power=.98, classification accuracy=.97.

*Symptoms of Complicated Bereavement*

The complicated bereavement symptoms of the RMHAP included both the DSM5 criteria of persistent complicated bereavement disorder (PCBD) and the draft ICD-11 definition of prolonged grief disorder (PGD). In our previous focus groups amongst West Papuan refugees in Port Moresby, participants endorsed the relevance of all the listed items, reporting that they aligned closely with the indigenous construct of loss and extreme grief referred to as “Duka Cita” in the lingua franca of the community, Bahasa Indonesian. Participants also suggested additional symptoms regarded as specific to experiences of grief related to conditions of persecution and displacement relevant to their context and culture. These included feelings of confusion, alienation, diminished sense of identity, and difficulties in planning for the future. In our Port Moresby study, complicated bereavement symptoms showed a high degree of stability over a six-month period as indicated by a non-significant change in mean symptom scores from baseline to follow-up (time1 to time2 difference=6.5, *P*=0.398).

In our present survey in Kiunga, respondents were required to meet the DSM5 entry criterion of experiencing at least one death or loss involving a family member and/or close friend in the past 12 months in order to proceed to the symptom criteria for complicated bereavement. In recording symptoms, we did not apply skip rules, that is, the inquiry was made into all items. The item pool in the present sample yielded a high level of internal reliability (Cronbach’s alpha=0.94).

*PTSD symptoms*

The same procedure was used to record PTSD symptoms which included both DSM5 and ICD-11 items. The wording of these symptoms was adapted to the culture and context following our qualitative work using focus groups in our prior Port Moresby study (American Psychiatric Association, 2013). Qualitative and psychometric data testing the reliability and validity of the measure have been reported in previous studies (Tay et al., 2015b). There was a high level of internal reliability for the pool of PTSD symptoms in the Kiunga sample (Cronbach’s alpha 0.97). The stability of symptoms was indicated by the absence of change in mean PTSD symptom scores over a six-month period in the Port Moresby study (time1 to time2 difference=.72, *P*=0.351).

For entry to the PTSD symptom module, respondents had to report at least one DSM5 defined traumatic event (for example, serious injuries to self, witnessing deaths/atrocities, sexual violence).

*Traumatic Loss Events*

Traumatic losses were assessed using five items derived from extensive community consultations detailing the broad types of relevant conflict-related events during the Indonesian occupation. A traumatic loss event count was generated by adding all endorsed items (scored 1) for each participant.

*Post-migration living difficulties (PMLDs)*

An inventory based on the Humanitarian Emergency Settings Perceived Needs (HESPER) scale (Semrau et al., 2012) was used to assess the prevalence of common forms of ongoing stressors in the community. We drew on information relevant to the local culture and context to modify this list based on an extensive series of consultations (based on focus groups, individual informant interviews) undertaken with members of the Kiunga community. The measure comprises 26 items rated on a four-point scale (0=not a problem, 1=a bit of a problem, 2=a moderately serious problem, 3=a very serious problem). For the present analysis, we generated a mean postmigration living difficulty (PMLD) score (mean range: 0.27—4.06) by adding up individual item scores (0-3).

*Adaptive Stress Index (ASI)*

The original Adaptive Stress Index (ASI) based on the ADAPT model was designed and applied in the earlier study amongst West Papuans residing in Port Moresby. The aim was to record the stress of adaptation arising from the undermining of the five psychosocial systems identified by the ADAPT model, contextualized to the history and ongoing conditions experienced by refugees from West Papua. The initial pool comprised 62 items each rated on a four-point severity scale (0=none; 1=moderately; 2=very strongly; 3=extremely) (Tay et al., 2015b). Details of the sequence of qualitative and quantitative steps pursued to develop and test the cultural and contextual validity of the measure and its psychometric properties have been provided previously (Tay et al., 2015a). The qualitative component involved an iterative process involving individual and focus group interviews. The information and feedback recorded during these inquiries allowed us to make iterative modifications to the content and expression of items. High levels of internal consistency were found for the full pool of items (α=0.97) and for each of the five ADAPT subscales (safety/security: α=0.89; attachments: α=0.92; access to justice: α=0.94; roles/identities: α=0.87; existential meaning: α=0.88). Confirmatory factor analysis supported a five-factor structure corresponding to the ADAPT pillars.

***Procedure***

Interviews were conducted by a field team in Bahasa Indonesian, English, and *Tok Pisin* (the pigeon English that is widely spoken throughout PNG), depending on the preference of the participant. The field interviewers were drawn from the West Papuan community and the team was managed by a West Papuan refugee (MK) who has worked with our research team on similar projects for several years.

***Field team and training***

Field interviewers were selected based on their fluency in the relevant languages; their status of trustworthiness in the community; and their competence and commitment to the project. The team included an equal number of men and women. Members of the team received three weeks’ intensive training from a clinical psychologist who speaks Bahasa Indonesian followed by three months of piloting the interview in the field. The training program, applied successfully in the preceding Port Moresby study, focused on interviewing techniques, identification of mental health issues amongst trauma survivors, role-play, and administration of the R-MHAP. We took active steps to minimize the risk of inadvertent disclosure of sensitive information revealed in interviews in the training by giving special focus to ethical issues of privacy and confidentiality in training and subsequent supervision.

A test of inter-rater reliability in the previous Port Moresby study achieved a 95% overall percentage agreement between field workers and professional personnel in assigning mental health diagnoses. Following training, the professional team conducted weekly supervision sessions for field workers via telephone or on-site for the full duration of the study period, extending from March to September 2016.

***Statistical analysis***

We applied latent class analysis (LCA) to identify subpopulations of West Papuans manifesting differential patterns of complicated bereavement and PTSD, including all symptoms of each category (each symptom scored 1 if present). The analysis was conducted on the whole sample including individuals who did not meet full criteria for either PTSD or PCBD. LCA is a person-centered classification method in which individuals are grouped according to specified characteristics. The classification is latent in that each person makes a proportional contribution to the classes that emerge based on their individual pattern of responses to the list of symptoms (Nylund et al., 2007). We tested sequential models (1—5 classes seriatim). Conventional indicators used included the Bayesian Information Criterion (BIC), sample size-adjusted Bayesian Information Criterion (SS-BIC), and the Akaike’s Information Criterion (AIC) (McGCutcheon, 1987, Collins and Lanza, 2009). Lower values of these indicators indicate a better fit in comparing successive latent class models. In addition, we applied the Vuong-Lo-Mendell-Rubin (VLMR) and the Lo-Mendell-Rubin (LMR) adjusted likelihood ratio tests, both of which compare the fit of a latent class model of n classes to one with n+1 classes (Nylund et al., 2007). In judging the best-fitting model, we considered the principle of parsimony (fewer classes are preferred over many), the degree of class separation, homogeneity of posterior probabilities within classes, and the interpretability of the classes yielded (Collins and Lanza, 2009). A conventional set of criteria have been proposed to categorize posterior probabilities (Burstein et al., 2012). According to that classification, conditional probabilities of 0.60 or above indicate a high probability of endorsing a symptom; values falling between 0.59 and 0.15, a moderate probability; and a value of 0.14 or less, a low probability.

After identifying the best-fitting model, we examined for associations between class membership and relevant predictors (traumatic events, post-migration living difficulties, psychosocial disruptions according to the ASI) using multinomial logistic regression analysis. The low symptom class was the reference category in all comparisons. Restriction in numbers precluded direct comparisons amongst the three derived morbid classes.

Tests of collinearity using the variance inflation factor (VIF), tolerance and inspection of eigenvalues indicated substantial covariance across the three predictor variables, as would be expected given that they overlapped in identifying related forms of adversity. We, therefore, tested each of the three substantive predictors in separate regression models, including sociodemographic variables (gender, marital status, employment) in each analysis as covariates. Analyses were performed in STATA version 13 and Mplus Version 7.